Calculation of interatomic forces and optimization of molecular geometry with auxiliary-field quantum Monte Carlo: Supplementary Information

Mario Motta[1] and Shiwei Zhang[2]

[1] *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA*

[2] *Department of Physics, College of William and Mary, Williamsburg, Virginia 23187-8795, USA*

The starting point for the evaluation of forces is the general form of the ground-state expectation value of the Hamiltonian operator,

$$E(\mathbf{R}) = \frac{\int d\Phi d\Psi\, \rho_{m,n}(\Phi,\Psi)\, W(\Phi,\Psi;\mathbf{R})\, E_{loc}(\Phi,\Psi;\mathbf{R})}{\int d\Phi d\Psi\, \rho_{m,n}(\Phi,\Psi)\, W(\Phi,\Psi;\mathbf{R})}\,. \tag{1}$$

The Slater determinant $|\Psi\rangle = \prod_\sigma \prod_{l=1}^{N_\sigma} \psi^\dagger_{l\sigma}|\emptyset\rangle$ has N_σ ($\sigma =\uparrow$ or $\downarrow$) fermions occupying orbitals $|\psi_{l\sigma}\rangle \equiv \psi^\dagger_{l\sigma}|\emptyset\rangle = \sum_i (C_{\Psi,\sigma})_{il}\,|\chi_i\rangle$, where $\{|\chi_i\rangle, i = 1\dots M\}$ are the basis functions and $C_{\Psi,\sigma}$ is a $M \times N_\sigma$ matrix that contains the orbital coefficients. The Slater determinant $|\Phi\rangle$ is similarly defined in terms of orbitals $|\phi_{l\sigma}\rangle$ and matrix $C_{\Phi,\sigma}$. In Eq. (1), the overlap is given by

$$\begin{aligned} W(\Phi,\Psi;\mathbf{R}) &\equiv \langle\Phi|\Psi\rangle \\ &= \prod_\sigma \det\left(C^\dagger_{\Phi,\sigma} S(\mathbf{R}) C_{\Psi,\sigma}\right) \equiv \prod_\sigma \det\left(\Omega_\sigma(\mathbf{R})\right)\,, \end{aligned} \tag{2}$$

where $S(\mathbf{R})_{ij} = \langle\chi_i|\chi_j\rangle$ is the overlap matrix of the AO basis set. The local energy is given by

$$E_{loc}(\Phi,\Psi;\mathbf{R}) = H_0(\mathbf{R}) + \sum_\sigma \mathrm{Tr}[h(\mathbf{R})G_\sigma(\mathbf{R})] + \sum_\gamma \Big(\sum_\sigma \mathrm{Tr}[L^\gamma(\mathbf{R})G_\sigma(\mathbf{R})]\Big)^2 - \sum_{\gamma\sigma} \mathrm{Tr}[L^\gamma(\mathbf{R})G_\sigma(\mathbf{R})L^\gamma(\mathbf{R})G_\sigma(\mathbf{R})]\,, \tag{3}$$

where $H_0(\mathbf{R})$ is a constant from the internuclear repulsion, the matrix $h(\mathbf{R})$ is the one-body part of the Hamiltonian in the AO basis ($h(\mathbf{R})_{ij} = \langle\chi_i|h_1(\mathbf{R})|\chi_j\rangle$), the matrices $L^\gamma(\mathbf{R})$ arise from decomposing the two-body part of the Hamiltonian (e.g., via the Cholesky or density-fitting decomposition [1, 2], $(ik|jl) = \sum_\gamma L^\gamma(\mathbf{R})_{ik} L^\gamma(\mathbf{R})_{jl}$), and the (spin-dependent) one-electron Green's function is

$$G_\sigma(\mathbf{R}) = C_{\Psi,\sigma}\Omega_\sigma(\mathbf{R})^{-1} C^\dagger_{\Phi,\sigma}\,, \tag{4}$$

which obeys the idempotence relation

$$G_\sigma(\mathbf{R}) = G_\sigma(\mathbf{R}) S_\sigma(\mathbf{R}) G_\sigma(\mathbf{R})\,. \tag{5}$$

The gradient of W can be obtained from Jacobi's formula for the derivative of the determinant,

$$\frac{\partial_{R_\alpha} W(\Phi,\Psi;\mathbf{R})}{W(\Phi,\Psi;\mathbf{R})} = \sum_\sigma \mathrm{Tr}\left(\partial_{R_\alpha} S(\mathbf{R})\, G_\sigma(\mathbf{R})\right)\,. \tag{6}$$

The gradient of $E_{loc}(\Phi,\Psi;\mathbf{R})$ is readily computed by recalling the cyclicity of the trace, and that differentiation of (4) yields

$$\partial_{R_\alpha} G_\sigma(\mathbf{R}) = -G_\sigma(\mathbf{R})\, \partial_{R_\alpha} S_\sigma(\mathbf{R})\, G_\sigma(\mathbf{R})\,, \tag{7}$$

complying with the idempotence relation (5) [3]. The gradient $\partial_{R_\alpha} G_\sigma(\mathbf{R})$ incorporates Pulay's corrections in the framework of AFQMC, which are important for an accurate evaluation of forces when atom-centered Gaussian orbitals are used as basis for the one-electron Hilbert space. The presence of Pulay's corrections is reminiscent of the covariant derivatives that appear in fiber bundle theory and differential geometry [4]. This reminiscence is not accidental: the purpose of corrective terms in covariant derivatives is to preserve some properties of the space on which they act. For example, in gauge theories corrections preserve the gradient of the wavefunction under gauge transformations and, on Riemann manifolds, connections preserve parallelism. In this framework, Pulay's correction preserve the idempotency (5) of the one-particle reduced Green's function under changes in molecular geometries.

The gradient of $E_{loc}(\Phi,\Psi;\mathbf{R})$ thus reads

$$\begin{aligned}\partial_{R_\alpha} E_{loc}(\Phi,\Psi;\mathbf{R}) &= \partial_{R_\alpha} H_0(\mathbf{R}) + \sum_\sigma \mathrm{Tr}[\partial_{R_\alpha} h(\mathbf{R})\, G_\sigma(\mathbf{R})] + \sum_\sigma \mathrm{Tr}[h(\mathbf{R})\, \partial_{R_\alpha} G_\sigma(\mathbf{R})] \\ &+2 \sum_{\gamma,\sigma\sigma'} \mathrm{Tr}[L^\gamma(\mathbf{R}) G_\sigma(\mathbf{R})]\, \mathrm{Tr}[\partial_{R_\alpha} L^\gamma(\mathbf{R})\, G_{\sigma'}(\mathbf{R})] - 2 \sum_{\gamma,\sigma} \mathrm{Tr}[\partial_{R_\alpha} L^\gamma(\mathbf{R})\, G_\sigma(\mathbf{R}) L^\gamma(\mathbf{R}) G_\sigma(\mathbf{R})] + \\ &+2 \sum_{\gamma,\sigma\sigma'} \mathrm{Tr}[L^\gamma(\mathbf{R}) G_\sigma(\mathbf{R})]\, \mathrm{Tr}[L^\gamma(\mathbf{R})\, \partial_{R_\alpha} G_{\sigma'}(\mathbf{R})] - 2 \sum_{\gamma,\sigma} \mathrm{Tr}[L^\gamma(\mathbf{R}) G_\sigma(\mathbf{R}) L^\gamma(\mathbf{R})\, \partial_{R_\alpha} G_\sigma(\mathbf{R})]\,. \end{aligned} \tag{8}$$

Computing the local energy requires $\mathcal{O}(N_\gamma(M^3 + M + M^2))$ operations, while computing all components of the force requires $\mathcal{O}(N_\gamma(3M^3 + M + M^2 + 4N_{R_\alpha}M^2))$ operations, with N_{R_α} being the number of force components. Thus the cost for computing all force components is about a factor $(3 + 4\frac{N_p}{M})$ of that of the local energy. The ratio is of course bounded by 7 and should be approximately 3 in most situations, which was confirmed in our studies. It should be possible to exploit the structure of the components of L and possibly regroup terms in Eq. (8) to further speed up the computation. The AFQMC estimator of the force is then given by

$$F^\alpha(\mathbf{R}) = \frac{\int d\Phi d\Psi\, \rho_{m,n}(\Phi,\Psi)\, W(\Phi,\Psi;\mathbf{R})\, \left(\partial_{R_\alpha} E_{loc}(\Phi,\Psi;\mathbf{R}) + \frac{\partial_{R_\alpha} W(\Phi,\Psi;\mathbf{R})}{W(\Phi,\Psi;\mathbf{R})}\left(E_{loc}(\Phi,\Psi;\mathbf{R}) - E(\mathbf{R})\right)\right)}{\int d\Phi d\Psi\, \rho_{m,n}(\Phi,\Psi)\, W(\Phi,\Psi;\mathbf{R})}\,. \tag{9}$$

When $m = n = 0$, $\rho_{m,n}(\Phi,\Psi) = \delta(\Phi - \Psi_T)\delta(\Psi - \Psi_T)$, so that $E(\mathbf{R}) = E_{loc}(\Psi_T,\Psi_T;\mathbf{R})$ and the estimator reduces to the familiar Hartree-Fock expression [3].

In this work, the derivatives $\partial_{R_\alpha} S$, $\partial_{R_\alpha} h$ and $\partial_{R_\alpha} L^\gamma$ were all obtained by numerical differentiation. A finite difference step size of 10^{-4} a.u. was used, and it was verified that using a smaller step does not change the results within statistical error bars.

[1] W. Purwanto, H. Krakauer, Y. Virgus, and S. Zhang, "Assessing weak hydrogen binding on Ca centers: an accurate many-body study with large basis sets," *J. Chem. Phys.*, vol. 135, no. 16, p. 164105, 2011.

[2] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. McClain, E. R.Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, "The python-based simulations of chemistry framework (pyscf)," *WIREs Comput. Mol. Sci.*

[3] P. Pulay, "Analytical derivatives, forces, force constants, molecular geometries, and related response properties in electronic structure theory," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 4, no. 3, pp. 169–181, 2014.

[4] T. U. Helgaker, "Hamiltonian expansion in geometrical distortions," in *Geometrical Derivatives of Energy Surfaces and Molecular Properties* (J. S. Poul Jorgensen, ed.), ch. 1, NATO Science Series C, 1986.